

## **RAW SEQUENCE LISTING**

**The Biotechnology Systems Branch of the Scientific and Technical  
Information Center (STIC) no errors detected.**

Application Serial Number: 10/565,233  
Source: IFWP  
Date Processed by STIC: 1-27-06

# ***ENTERED***



IFWP

## RAW SEQUENCE LISTING

DATE: 01/27/2006

PATENT APPLICATION: US/10/565,233

TIME: 09:42:23

Input Set : A:\Sequence Listing.txt

Output Set: N:\CRF4\01272006\J565233.raw

```

3 <110> APPLICANT: Washington State University Research Foundation
4     Rodney, Croteau
5     Stefan, Jennewein
6     Robert, Long
8 <120> TITLE OF INVENTION: P450 OXYGENASES AND METHODS OF USE
10 <130> FILE REFERENCE: 4630-66380-05
C--> 12 <140> CURRENT APPLICATION NUMBER: US/10/565,233
C--> 12 <141> CURRENT FILING DATE: 2006-01-18
12 <150> PRIOR APPLICATION NUMBER: PCT/US2004/023656
13 <151> PRIOR FILING DATE: 2004-07-21
15 <150> PRIOR APPLICATION NUMBER: US 60/489,597
16 <151> PRIOR FILING DATE: 2003-07-22
18 <160> NUMBER OF SEQ ID NOS: 41
20 <170> SOFTWARE: PatentIn version 3.2
22 <210> SEQ ID NO: 1
23 <211> LENGTH: 1688
24 <212> TYPE: DNA
25 <213> ORGANISM: Taxus cuspidata
28 <220> FEATURE:
29 <221> NAME/KEY: CDS
30 <222> LOCATION: (40)..(1548)
31 <223> OTHER INFORMATION: coding sequence
33 <400> SEQUENCE: 1
34 cggcaccagg tttctgctc ctgcttttct tcttccaaa atg gac gcc ctg tat      54
35                                     Met Asp Ala Leu Tyr
36                                     1       5
38 aag agc aca gtt gca aaa ttt aat gag gtc aca cag ctg gac tgt tcc      102
39 Lys Ser Thr Val Ala Lys Phe Asn Glu Val Thr Gln Leu Asp Cys Ser
40       10       15       20
42 act gaa tct ttt tcc att gcc ctc tca gct att gct ggt att ctt ctg      150
43 Thr Glu Ser Phe Ser Ile Ala Leu Ser Ala Ile Ala Gly Ile Leu Leu
44       25       30       35
46 ctt ctc ctg ctc ttc cgt tct aaa cgc cac tcc tcc ctt aaa ctt cct      198
47 Leu Leu Leu Leu Phe Arg Ser Lys Arg His Ser Ser Leu Lys Leu Pro
48       40       45       50
50 cct ggg aaa tta ggc atc cct ttc att ggc gag tcg ttt atc ttc ctg      246
51 Pro Gly Lys Leu Gly Ile Pro Phe Ile Gly Glu Ser Phe Ile Phe Leu
52       55       60       65
54 agg gct ctt cga tcg aac tcg ctg gag caa ttt ttt gac gag aga gtg      294
55 Arg Ala Leu Arg Ser Asn Ser Leu Glu Gln Phe Phe Asp Glu Arg Val
56 70       75       80       85
58 aag aaa ttc ggc ctc gtg ttc aag acc tcc ttg att ggg cat ccc aca      342
59 Lys Lys Phe Gly Leu Val Phe Lys Thr Ser Leu Ile Gly His Pro Thr

```

## RAW SEQUENCE LISTING

DATE: 01/27/2006

PATENT APPLICATION: US/10/565,233

TIME: 09:42:23

Input Set : A:\Sequence Listing.txt

Output Set: N:\CRF4\01272006\J565233.raw

60	90	95	100	
62 gta gta ctc tgc ggc cct gcg gga aac cgg ctt att ctg tcc aac gag				390
63 Val Val Leu Cys Gly Pro Ala Gly Asn Arg Leu Ile Leu Ser Asn Glu				
64	105	110	115	
66 gag aag ctg gtg cag atg tcg tgg ccc gct caa ttt atg aag ctc atg				438
67 Glu Lys Leu Val Gln Met Ser Trp Pro Ala Gln Phe Met Lys Leu Met				
68	120	125	130	
70 ggg gag aat tcc gtt gcc acc agg agg ggt gaa gac cat ata gtt atg				486
71 Gly Glu Asn Ser Val Ala Thr Arg Arg Gly Glu Asp His Ile Val Met				
72	135	140	145	
74 cgc tct gct ctt gca ggt ttt ttc ggc cct ggt gcg ctg cag agt tac				534
75 Arg Ser Ala Leu Ala Gly Phe Phe Gly Pro Gly Ala Leu Gln Ser Tyr				
76 150	155	160	165	
78 att ggt aaa atg aat aca gag atc cag agt cat atc aac gaa aaa tgg				582
79 Ile Gly Lys Met Asn Thr Glu Ile Gln Ser His Ile Asn Glu Lys Trp				
80	170	175	180	
82 aag gga aaa gat gag gtg aat gta ctt cct ttg gta aga gag ctc gtc				630
83 Lys Gly Lys Asp Glu Val Asn Val Leu Pro Leu Val Arg Glu Leu Val				
84	185	190	195	
86 ttc aac att tcg gcc atc ttg ttt ttc aac ata tat gat aag cag gaa				678
87 Phe Asn Ile Ser Ala Ile Leu Phe Phe Asn Ile Tyr Asp Lys Gln Glu				
88	200	205	210	
90 cag gat cgt ctg cat aag ctt ttg gaa act att ctg gtc gga agt ttt				726
91 Gln Asp Arg Leu His Lys Leu Leu Glu Thr Ile Leu Val Gly Ser Phe				
92	215	220	225	
94 gct ctt ccg att gac ttg ccc gga ttt ggt ttc cat aga gca ctc cag				774
95 Ala Leu Pro Ile Asp Leu Pro Gly Phe Gly Phe His Arg Ala Leu Gln				
96 230	235	240	245	
98 gga cgg gcc aag ctc aac aaa att atg ctg tct tta att aaa aag aga				822
99 Gly Arg Ala Lys Leu Asn Lys Ile Met Leu Ser Leu Ile Lys Lys Arg				
100	250	255	260	
102 aaa gaa gat ctg cag tct gga tcg gca aca gcc acg cag gat ctg ctc				870
103 Lys Glu Asp Leu Gln Ser Gly Ser Ala Thr Ala Thr Gln Asp Leu Leu				
104	265	270	275	
106 tct gtt ttg ctc act ttc aga gat gac aaa ggg act cca ctc acc aat				918
107 Ser Val Leu Leu Thr Phe Arg Asp Lys Gly Thr Pro Leu Thr Asn				
108	280	285	290	
110 gat gag ata ctc gac aac ttt tct tct ctg ctc cat gcc tcc tat gac				966
111 Asp Glu Ile Leu Asp Asn Phe Ser Ser Leu Leu His Ala Ser Tyr Asp				
112	295	300	305	
114 acc acc act tcg cca atg gct ttg att ttc aag ctc ttg tct tcc aat				1014
115 Thr Thr Thr Ser Pro Met Ala Leu Ile Phe Lys Leu Leu Ser Ser Asn				
116 310	315	320	325	
118 cca gaa tgc tat caa aaa gta gtt caa gag caa ttg gag ata ctt tcc				1062
119 Pro Glu Cys Tyr Gln Lys Val Val Gln Glu Gln Leu Glu Ile Leu Ser				
120	330	335	340	
122 aac aaa gag gag ggc gaa gaa atc aca tgg aag gat ctc aaa gcc atg				1110
123 Asn Lys Glu Glu Gly Glu Glu Ile Thr Trp Lys Asp Leu Lys Ala Met				
124	345	350	355	

## RAW SEQUENCE LISTING

DATE: 01/27/2006

PATENT APPLICATION: US/10/565,233

TIME: 09:42:23

Input Set : A:\Sequence Listing.txt

Output Set: N:\CRF4\01272006\J565233.raw

```

126 aaa tac aca tgg caa gta gct cag gaa acg ctg cgg atg ttt cct cca      1158
127 Lys Tyr Thr Trp Gln Val Ala Gln Glu Thr Leu Arg Met Phe Pro Pro
128          360                      365                      370
130 gtt ttc gga aca ttt cgc aag gcc atc act gac att cag tat gat ggt      1206
131 Val Phe Gly Thr Phe Arg Lys Ala Ile Thr Asp Ile Gln Tyr Asp Gly
132          375                      380                      385
134 tac aca att cca aaa ggg tgg aag ctg ttg tgg aca act tac agt aca      1254
135 Tyr Thr Ile Pro Lys Gly Trp Lys Leu Leu Trp Thr Thr Tyr Ser Thr
136 390          395                      400                      405
138 cat ccc aag gac ttg tat ttc aat gaa cca gag aaa ttc atg cct tca      1302
139 His Pro Lys Asp Leu Tyr Phe Asn Glu Pro Glu Lys Phe Met Pro Ser
140          410                      415                      420
142 aga ttc gat cag gaa gga aag cat gta gct cct tac aca ttt ttg ccc      1350
143 Arg Phe Asp Gln Glu Gly Lys His Val Ala Pro Tyr Thr Phe Leu Pro
144          425                      430                      435
146 ttc ggt gga ggc caa cgg tca tgt gtg gga tgg gaa ttt tca aag atg      1398
147 Phe Gly Gly Gly Gln Arg Ser Cys Val Gly Trp Glu Phe Ser Lys Met
148          440                      445                      450
150 gag ata tta cta ttc gtt cat cat ttt gtc aaa act ttt agc agc tac      1446
151 Glu Ile Leu Leu Phe Val His Phe Val Lys Thr Phe Ser Ser Tyr
152          455                      460                      465
154 acc cca gtt gat ccc gac gaa aaa ata tca ggg gat cca ctc cct cct      1494
155 Thr Pro Val Asp Pro Asp Glu Lys Ile Ser Gly Asp Pro Leu Pro Pro
156 470          475                      480                      485
158 ctt cct tcc aag gga ttt tcc att aaa ctg ttt ccc gag acc ata gtc      1542
159 Leu Pro Ser Lys Gly Phe Ser Ile Lys Leu Phe Pro Glu Thr Ile Val
160          490                      495                      500
162 aat tga aggagaaaac cacagtgcag aactgctatt cttgaatcct cgctcaagaa      1598
163 Asn
166 taatacaaac atgcatcacc aacaatgttt atgcactcaa tgcaaattaa cagtgtgtca      1658
168 gcattgacag tcaaaaaaaaa aaaaaaaaaa      1688
171 <210> SEQ ID NO: 2
172 <211> LENGTH: 502
173 <212> TYPE: PRT
174 <213> ORGANISM: Taxus cuspidata
176 <400> SEQUENCE: 2
178 Met Asp Ala Leu Tyr Lys Ser Thr Val Ala Lys Phe Asn Glu Val Thr
179 1          5          10          15
182 Gln Leu Asp Cys Ser Thr Glu Ser Phe Ser Ile Ala Leu Ser Ala Ile
183          20          25          30
186 Ala Gly Ile Leu Leu Leu Leu Leu Leu Phe Arg Ser Lys Arg His Ser
187          35          40          45
190 Ser Leu Lys Leu Pro Pro Gly Lys Leu Gly Ile Pro Phe Ile Gly Glu
191          50          55          60
194 Ser Phe Ile Phe Leu Arg Ala Leu Arg Ser Asn Ser Leu Glu Gln Phe
195 65          70          75          80
198 Phe Asp Glu Arg Val Lys Lys Phe Gly Leu Val Phe Lys Thr Ser Leu
199          85          90          95
202 Ile Gly His Pro Thr Val Val Leu Cys Gly Pro Ala Gly Asn Arg Leu

```

## RAW SEQUENCE LISTING

DATE: 01/27/2006

PATENT APPLICATION: US/10/565,233

TIME: 09:42:23

Input Set : A:\Sequence Listing.txt

Output Set: N:\CRF4\01272006\J565233.raw

203				100					105				110
206	Ile	Leu	Ser	Asn	Glu	Glu	Lys	Leu	Val	Gln	Met	Ser	Trp
207				115					120				125
210	Phe	Met	Lys	Leu	Met	Gly	Glu	Asn	Ser	Val	Ala	Thr	Arg
211				130					135				140
214	Asp	His	Ile	Val	Met	Arg	Ser	Ala	Leu	Ala	Gly	Phe	Phe
215	145										155		160
218	Ala	Leu	Gln	Ser	Tyr	Ile	Gly	Lys	Met	Asn	Thr	Glu	Ile
219					165					170			175
222	Ile	Asn	Glu	Lys	Trp	Lys	Gly	Lys	Asp	Glu	Val	Asn	Val
223					180					185			190
226	Val	Arg	Glu	Leu	Val	Phe	Asn	Ile	Ser	Ala	Ile	Leu	Phe
227					195					200			205
230	Tyr	Asp	Lys	Gln	Glu	Gln	Asp	Arg	Leu	His	Lys	Leu	Leu
231				210					215				220
234	Leu	Val	Gly	Ser	Phe	Ala	Leu	Pro	Ile	Asp	Leu	Pro	Gly
235	225					230					235		240
238	His	Arg	Ala	Leu	Gln	Gly	Arg	Ala	Lys	Leu	Asn	Lys	Ile
239					245					250			255
242	Leu	Ile	Lys	Lys	Arg	Lys	Glu	Asp	Leu	Gln	Ser	Gly	Ser
243					260					265			270
246	Thr	Gln	Asp	Leu	Leu	Ser	Val	Leu	Leu	Thr	Phe	Arg	Asp
247					275					280			285
250	Thr	Pro	Leu	Thr	Asn	Asp	Glu	Ile	Leu	Asp	Asn	Phe	Ser
251					290					295			300
254	His	Ala	Ser	Tyr	Asp	Thr	Thr	Thr	Ser	Pro	Met	Ala	Leu
255	305					310					315		320
258	Leu	Leu	Ser	Ser	Asn	Pro	Glu	Cys	Tyr	Gln	Lys	Val	Val
259					325					330			335
262	Leu	Glu	Ile	Leu	Ser	Asn	Lys	Glu	Glu	Gly	Glu	Glu	Ile
263					340					345			350
266	Asp	Leu	Lys	Ala	Met	Lys	Tyr	Thr	Trp	Gln	Val	Ala	Gln
267					355					360			365
270	Arg	Met	Phe	Pro	Pro	Val	Phe	Gly	Thr	Phe	Arg	Lys	Ala
271					370					375			380
274	Ile	Gln	Tyr	Asp	Gly	Tyr	Thr	Ile	Pro	Lys	Gly	Trp	Lys
275	385					390					395		400
278	Thr	Thr	Tyr	Ser	Thr	His	Pro	Lys	Asp	Leu	Tyr	Phe	Asn
279					405					410			415
282	Lys	Phe	Met	Pro	Ser	Arg	Phe	Asp	Gln	Glu	Gly	Lys	His
283					420					425			430
286	Tyr	Thr	Phe	Leu	Pro	Phe	Gly	Gly	Gly	Gln	Arg	Ser	Cys
287					435					440			445
290	Glu	Phe	Ser	Lys	Met	Glu	Ile	Leu	Leu	Phe	Val	His	His
291					450					455			460
294	Thr	Phe	Ser	Ser	Tyr	Thr	Pro	Val	Asp	Pro	Asp	Glu	Lys
295	465					470					475		480
298	Asp	Pro	Leu	Pro	Pro	Leu	Pro	Ser	Lys	Gly	Phe	Ser	Ile
299					485					490			495

## RAW SEQUENCE LISTING

DATE: 01/27/2006

PATENT APPLICATION: US/10/565,233

TIME: 09:42:23

Input Set : A:\Sequence Listing.txt

Output Set: N:\CRF4\01272006\J565233.raw

302 Pro Glu Thr Ile Val Asn

303 500

306 &lt;210&gt; SEQ ID NO: 3

307 &lt;211&gt; LENGTH: 1455

308 &lt;212&gt; TYPE: DNA

309 &lt;213&gt; ORGANISM: Taxus cuspidata

311 &lt;400&gt; SEQUENCE: 3

```

312 atggatacct tcattcagca cgagtcttcc ccacttcttc tttctcttac tctcgtctgtt      60
314 attcttggca caattcttct tttgatatta agtggtaaag agtacagatc ttctcgtaaa      120
316 cttccccctg gaaacatggg cttccctctc attggggaga ctatagcact tatatcagat      180
318 acacctcgga agtttatcga cgacagagtg aagaaattcg gcctgggttt caagacttcg      240
320 ctaattgggc atcccgcagt tgtaatatgc ggctcctccg caaaccgttt cctcctctcc      300
322 aacgaggaaa agctgggtgc gatgtctttg cccaacgcag tactgaaact cttggggcag      360
324 gattgcgtta tggggaaaac cggagtgagg catgggattg tacgtaccgc actagcccgc      420
326 gccttggggc cccaggcggt gcagaattat gtggccaaaa tgagttcaga gatcgaacac      480
328 catatcaacc aaaaatggaa ggggaaagat gaggtgaagg tgcttctctt gataagaagc      540
330 ctcgtcttct ccatttcaac cagcttggtt ttccgtataa acgatgagca ccaacagaag      600
332 cgacttcata atcttttggg aactgtagct atgggacttg tgagtattcc cctagacttt      660
334 ccaggaactc gttttcgtaa agcactttac gcgcggtcga agctcgatga aattatgtct      720
336 tctgtaatag aaaggagaag aagcgatctt cgttcaggag cagcttcaag cgaccaagat      780
338 ctactgtcgg tgttgggtcac cttcaaagat gaaagaggga attcattcgc agacaaggag      840
340 atactggata acttctcttt tctacttcac gccttatacg acaccacaat ttcaccactc      900
342 accttgatat ttaagctgct ctcctctagt cctgaatgct atgagaatat agctcaagag      960
344 cagctggaaa tacttggcaa taaaaaggat agagaggaaa tcagctggaa ggatctgaag     1020
346 gatatgaaat atacatggca agcagttcag gaaactttga ggatgttccc tccagtttat     1080
348 ggatatattc gcgaggcttt gacagatatt gactatgatg gctatacaat accaaaagga     1140
350 tggagaatat tatgttcacc tcatactacg catagtaaag aggagtattt cgatgagccg     1200
352 gaagaattca gaccttcaag attcgaggat caaggaaggc atgtggctcc ttacacattc     1260
354 ataccatttg gaggaggcct gcgcattctg gcaggctggg aatttgcaaa gatggagata     1320
356 ttactgttta tgcattcatt tgttaaaact ttcagtcact tcattccagt tgaccccaac     1380
358 gaaaagattt cgagagatcc actgcctccc atccctgtca aaggattttc cataaagcct     1440
360 tttcctagat cataa                                     1455

```

363 &lt;210&gt; SEQ ID NO: 4

364 &lt;211&gt; LENGTH: 484

365 &lt;212&gt; TYPE: PRT

366 &lt;213&gt; ORGANISM: Taxus cuspidata

368 &lt;400&gt; SEQUENCE: 4

```

370 Met Asp Thr Phe Ile Gln His Glu Ser Ser Pro Leu Leu Leu Ser Leu
371 1      5      10      15
374 Thr Leu Ala Val Ile Leu Gly Thr Ile Leu Leu Leu Ile Leu Ser Gly
375      20      25      30
378 Lys Gln Tyr Arg Ser Ser Arg Lys Leu Pro Pro Gly Asn Met Gly Phe
379      35      40      45
382 Pro Leu Ile Gly Glu Thr Ile Ala Leu Ile Ser Asp Thr Pro Arg Lys
383      50      55      60
386 Phe Ile Asp Asp Arg Val Lys Lys Phe Gly Leu Val Phe Lys Thr Ser
387 65      70      75      80
390 Leu Ile Gly His Pro Ala Val Val Ile Cys Gly Ser Ser Ala Asn Arg
391      85      90      95

```

RAW SEQUENCE LISTING ERROR SUMMARY  
PATENT APPLICATION: US/10/565,233

DATE: 01/27/2006  
TIME: 09:42:24

Input Set : A:\Sequence Listing.txt  
Output Set: N:\CRF4\01272006\J565233.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:9; N Pos. 68,510,1092,1100,1188 ✓  
Seq#:10; Xaa Pos. 23,366,371,402  
Seq#:13; N Pos. 1151  
Seq#:14; Xaa Pos. 384  
Seq#:25; N Pos. 6,9,14  
Seq#:26; N Pos. 6,9,12,15  
Seq#:27; N Pos. 3,4,5,6,9,12,18  
Seq#:28; N Pos. 12,15  
Seq#:29; N Pos. 10,11,12,15,18  
Seq#:30; N Pos. 6,7,8,9,12,13,14,15,18,21,27  
Seq#:31; N Pos. 3,9,12,18,21,24,30

## VERIFICATION SUMMARY

DATE: 01/27/2006

PATENT APPLICATION: US/10/565,233

TIME: 09:42:24

Input Set : A:\Sequence Listing.txt

Output Set: N:\CRF4\01272006\J565233.raw

L:12 M:270 C: Current Application Number differs, Replaced Current Application No

L:12 M:271 C: Current Filing Date differs, Replaced Current Filing Date

L:910 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:9 after pos.:60

L:924 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:9 after pos.:480

L:944 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:9 after pos.:1080

L:946 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:9 after pos.:1140

L:991 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:10 after pos.:16

L:1075 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:10 after pos.:352

L:1079 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:10 after pos.:368

L:1087 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:10 after pos.:400

L:1359 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:13 after pos.:1140

L:1479 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:14 after pos.:368

L:2872 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:25 after pos.:0

L:2905 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:26 after pos.:0

L:2938 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:27 after pos.:0

L:2961 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:28 after pos.:0

L:2989 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:29 after pos.:0

L:3027 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:30 after pos.:0

L:3075 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:31 after pos.:0